

次世代音声合成方式の研究開発とオープンソース戦略による社会貢献

受賞業績 隠れマルコフモデルに基づいた次世代音声合成方式の確立

徳田 恵一^{*1} 戸田 智基^{*2} 全 炳河^{*3} 山岸 順一^{*4} 大浦 圭一郎^{*1}

^{*1}名古屋工業大学 ^{*2}奈良先端科学技術大学院大学 ^{*3}Google ^{*4}エジンバラ大学/国立情報学研究所

音声インタフェースが社会に広く普及し、日々の生活の中で多くの人々に喜んで使われるためには、音声認識等の性能改善もさることながら、さまざまな場面や状況に応じて感情表現や強調等を含む音声を自在に生成できる必要があると考え、人間のように多様な音声を自在に生成することの可能な新しい音声合成の枠組み「HMM 音声合成」を提唱してきました。従来の代表的な音声合成方式が、音声単位を波形データとして接続することにより任意の文に対応する音声を生成しようとするのに対して、本方式では、そのパラメータ列を統計モデルの1つである隠れマルコフモデル (Hidden Markov Model : HMM) によりモデル化し、その統計量から音声を再合成するものです(図-1)。

本方式には、a) データに基づいたシステムの自動構築が可能、b) 言語依存性が低く、多言語への拡張が容易、c) 統計量に基づいて音声を生成するため、携帯電話等の限られた計算資源の中で動作可能、などの利点があります。さらに、従来方式にはない大きな利点として、HMM の統計量を表すモデルパラメータを適切に変換することにより、さまざまな話者の声質、感情表現を伴った音声等、多様な音声を容易に生成できるという特長があります。

研究成果はソフトウェアツールキット「HMM-based Speech Synthesis System (HTS)」および「hts_engine API」、「Open JTalk」、「SPTK」などの関連ソフトウェアツールとしてとりまとめ、公開してきました。ライセンスは、オープンソースソフトウェアの標準的なものの1つである修正 BSD を採用しました。コピーライト表示をすれば、改変、再配布、商用利用等が可能な比較的自由度の高いライセンス形態です。延べダウンロード数は 30,000 件を超えており、音声合成研究の基盤ソフトウェアのデファクトの1つの地位を確立していると思われます。本方式は、カーナビゲーション、携帯電話・スマートフォン等で広く用いら

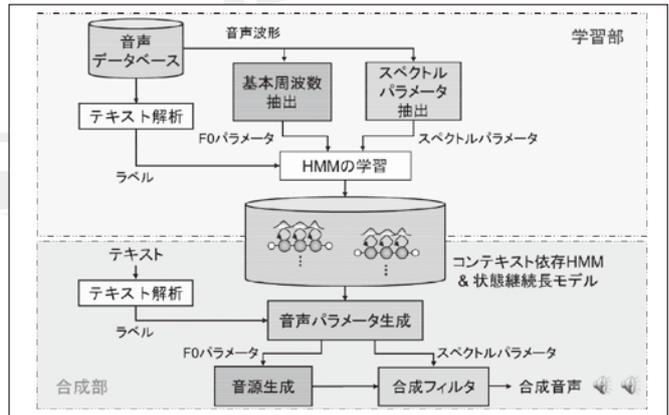


図-1 隠れマルコフモデル (Hidden Markov Model : HMM) を用いた新しい音声合成方式「HMM 音声合成」

れるようになってきていますが、これらのオープンソースソフトウェアによる波及効果も大きかったものと考えています。

最後になりましたが、荣誉ある賞をいただいたことを光栄に思うとともに関係者の皆様に感謝いたします。なお、本技術および関連ソフトウェアは、さまざまな研究機関の研究者の方々の協力により、発展してきたものです。関係各位に感謝申し上げます。(2013年5月21日受付)

徳田 恵一(正会員) tokuda@nitech.ac.jp

東京工業大学大学院博士課程修了。工学博士。現在、名古屋工業大学大学院教授。文部科学大臣表彰科学技術賞他受賞。現在、音声技術の普及にプレクスルーをもちます方法論について模索中。

戸田 智基(正会員) tomoki@is.naist.jp

平成 11 年名古屋大学工学部電気電子・情報工学科卒業。平成 15 年奈良先端科学技術大学院大学情報科学研究科博士課程修了。博士(工学)。現在、奈良先端科学技術大学院大学情報科学研究科准教授。

全 炳河(正会員) heigazen@google.com

平成 11 年鈴鹿工業高等専門学校電子情報工学科卒業。平成 18 年名古屋工業大学大学院情報工学専攻博士後期課程修了。博士(工学)。現在 Google Senior Research Scientist。

山岸 順一(正会員) jyamagis@nii.ac.jp

2006 年東京工業大学大学院博士課程修了。2007 ~ 2013 年英国エジンバラ大学 EPSRC Career Acceleration Fellow 他。2013 年より同大講師および国立情報学研究所(NII)准教授。工学博士。音声情報処理の研究に従事。

大浦 圭一郎(正会員) uratec@nitech.ac.jp

平成 22 年名古屋工業大学大学院工学研究科情報工学専攻博士後期課程修了。博士(工学)。現在、名古屋工業大学大学院特任助教。本会山下記念研究賞、日本音響学会独創研究奨励賞板倉記念等を受賞。